How well are DSM-5 diagnostic criteria for ASD represented in standardized diagnostic instruments?

Running head: DSM-5 in diagnostic instruments for ASD

Evers, Kris [*1,2,3], Maljaars, Jarymke [*1,2,3], Carrington, Sarah, J. [4], Carter, Alice, S. [5], Happé, Francesca [6], Steyaert, Jean [2,3,7], Leekam, Susan, R. [8], & Noens, Ilse [1,2]

[*] joint first authors

[1] Parenting and Special Education Research Unit, KU Leuven, Belgium; [2] Leuven Autism Research (LAuRes), KU Leuven, Belgium; [3] Child and Adolescent Psychiatry, Z.org UPC KU Leuven, Belgium; [4] School of Life and Health Sciences, Department of Psychology, Aston University, UK; [5] University of Massachusetts, Boston, USA; [6] Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK; [7] Center for Developmental Psychiatry, KU Leuven, Belgium; [8] Wales Autism Research Centre, School of Psychology, Cardiff University, Cardiff, UK

Word count (main text): 7,351 words

Correspondence concerning this article should be addressed to Kris Evers, Parenting and Special Education Research Unit, KU Leuven, Leopold Vanderkelenstraat 32, P.O. Box 3765, 3000 Leuven, Belgium**.** Tel: +32 16326018**,** E-mail: kris.evers@kuleuven.be

Abstract

Five years after the publication of DSM-5 in 2013, three widely-used diagnostic instruments have published algorithms designed to represent its (sub)criteria for Autism Spectrum Disorder (ASD) in children and adolescents. This study aimed to: (1) establish the content validity of these three DSM-5-adapted algorithms, and (2) identify problems with the operationalization of DSM-5 diagnostic criteria in measurable and observable behaviors. Algorithm items of the Autism Diagnostic Observation Schedule - Second Edition (ADOS-2), Developmental, Dimensional and Diagnostic Interview (3di) and Diagnostic Interview for Social and Communication Disorders - 11th edition (DISCO-11) were mapped onto DSM-5 sub-criteria. The development and decision-making rules integrated in their algorithms were then compared with DSM-5. Results demonstrated significant variability in the number and nature of sub-criteria covered by the ADOS-2, 3di and DISCO-11. In addition to differences in the development of algorithms and cut-off scores, instruments also differed in the extent to which they follow DSM-5 decision-making rules for diagnostic classification. We conclude that such differences in interpretation of DSM-5 criteria provide a challenge for symptom operationalization which will be most effectively overcome by consensus, testing and reformulation.

Keywords

Autism Spectrum Disorder; classification; content validity; diagnostic outcome

Acknowledgments

The authors thank Wouter De la Marche (child psychiatrist at OPZ Geel, Belgium) and Annelies de Bildt (research psychologist at Accare, the Netherlands) for their contribution to the coding process as members of the expert panel. This research was supported by a postdoctoral fellowship from the Fund for Scientific Research Flanders awarded to KE (12L6916N) and by the KU Leuven (Leuvens Universiteitsfonds 'Opening the Future').

The diagnostic assessment of Autism Spectrum Disorder (ASD) in children is a complex process, in which information is gathered from parents (or caregivers) about the child's developmental history and current level of functioning, together with first-hand observations by an experienced clinician [1–4]. Standardized semi-structured observation instruments and parental interviews are now widely used in this information-gathering process [1]. The narrow use of instruments, such as in only administering the algorithm items or focusing solely on the algorithm's outcomes for the purpose of diagnosis, should never be used to decide diagnostic classification, but instead diagnostic classification should rely on the integration of different sources of information, including a parental interview and a child observation and from different contexts [2, 5, 6]. Nevertheless, in a research context, clinical diagnoses of participants' are sometimes *validated* using a semi-structured observation instrument and/or a parental interview, and sometimes participants who do not meet the threshold are even excluded from the research sample, which could lead to a biased understanding. If we are consistently excluding individuals from research based on one particular sub-criterion (due to the fact that the instruments do not adequately measure it), we may not be best representing individuals with difficulties in that particular area. It is therefore important to study to what extent behaviors described by the DSM-5 criteria are represented in diagnostic assessment instruments for ASD, as well as the procedures by which a classification according to these DSM-5 criteria could be implemented. Gaining insight into the content validity of the algorithms can help clinicians understand why an individual meets the threshold on a specific instrument (or fails to do so), so that they can seek converging (or diverging) information of other sources. In this way, our study could be important in both supporting clinicians' decision-making processes and in facilitating parity of research samples recruited according to the DSM-5 criteria.

Autism Spectrum Disorder in DSM-5

The latest edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [7] includes significant changes to the diagnostic criteria for ASD. While DSM-IV-TR [8] delineated five different sub-classifications, DSM-5 abandoned those sub-classifications in favor of one single classification, Autism Spectrum Disorder (ASD). Additional changes were related to the diagnostic criteria. Instead of a triad of impairments, DSM-5 characterized ASD by deficits in two core domains: (1) impairments in social interaction and social communication, and (2) repetitive and restricted patterns of activity, behaviors and interests (RRBIs). More specifically, to meet DSM-5 criteria for ASD, individuals are required to meet all three sub-criteria within the social interaction and social communication domain, and two out of four of the sub-criteria within the RRBI

domain (for more details, see Appendix A). The latter rule gave greater significance to RRBIs; in DSM-IV-TR, only one of four RRBI sub-criteria had to be met. In addition, the number of possible combinations of sub-criteria that would qualify for an ASD diagnosis was limited from 2,027 for a DSM-IV-TR diagnosis to 11 possible combinations for a DSM-5 diagnosis [9]. Furthermore, sensory problems were added as a new symptom within the RRBI domain, and language problems were removed from the core ASD symptoms and considered instead as co-occurring difficulties (like intellectual disability) that can be indicated with a *specifier* in order to describe an individual's profile. Finally, DSM-5 stipulates levels of severity for both domains of impairment based on the required level of support.

Such a change in diagnostic criteria could significantly alter the characterization of autism with consequences for the number of individuals being diagnosed. Although DSM-5 explicitly states that individuals previously diagnosed with Autistic Disorder or Asperger's Disorder should qualify for a DSM-5 diagnosis of ASD, meta-analyses and literature studies suggest that a significant proportion of individuals who met DSM-IV-TR criteria will fail to meet DSM-5 criteria for ASD, especially those with a diagnosis of PDD-NOS or Asperger's Disorder [10–12].

Aims of the current study

DSM-5 was published in 2013 [7]. Recently, the International Classification of Diseases (ICD) has also published its novel guideline ICD-11 [13], paralleling DSM-5. Given that some authors have suggested that application of the new DSM-5 criteria can result in a shift and a decrease of ASD diagnoses (for a review, see [11]) the aim of the study was to document the effect of DSM-5 changes on existing diagnostic instruments that have been designed to guide diagnostic judgements. Specifically, the purpose of the paper is to systematically identify the way that these instruments operationally define diagnostic criteria and sub-criteria and if there is consistency between these instruments in the way that behavior is operationalized to match DSM-5 criteria. This paper is not aimed at evaluating the correctness of diagnostic classifications (as empirical studies of psychometric properties do), but this paper aims at the characterization of ASD behaviors and the consistency in the way they are operationalized. This operationalization or content validity also contributes to the clinical utility of an instrument, as it is crucial that clinicians/researchers gain insight in which sub-criteria are covered by the instruments, and how the algorithm is developed. Such insight can help clinicians to understand and analyze why an individual meets the threshold on a specific instrument (or fails to do so) and to seek evidence related to the not-covered criteria.

Three diagnostic instruments have developed new algorithms, specifically designed to measure DSM-5 criteria characteristics, but it is not yet clear whether these three instruments cover DSM-5 symptoms to the same extent and whether different procedures used by the three algorithms can lead to different diagnostic outcomes. As part of the development of some of the instruments, the specific instrument items were mapped onto DSM-5 criteria [for ADOS-2, see 14, 15, for DISCO-11, see 16], which is also presented in the Results section. The revised DSM-5 adapted algorithms of these instruments (or preliminary versions of them) demonstrated good psychometric properties [17–22, for a systematic review of psychometric characteristics of instruments available for preschoolers, see 23]. Previously, Huerta and colleagues [14] have studied the content validity of the ADOS-2, and concluded that the instrument did not cover all sub-criteria for ASD, but the DSM-5 algorithms of different instruments were not yet compared directly in terms of content validity. Therefore, the first goal of the study was to establish the content validity of these three DSM-5-adapted algorithms. The second was to evaluate the clarity of the DSM-5 criteria themselves and identify possible pitfalls when operationalizing the DSM-5 (sub-)criteria into concrete measurable behaviors. In this way, we hope to guide future improvements in diagnostic instruments and classification systems.

Method

Procedure

Selection of instruments. Autism-specific diagnostic interviews and observation schedules for children and adolescents with a wide age range were selected from the guidelines for diagnosis of autism developed by the National Institute for Health and Care Excellence [5], excluding screening instruments and questionnaires. Only instruments with newly developed DSM-5 based scoring principles were included, yielding the following three instruments: the Autism Diagnostic Observation Schedule - Second Edition (ADOS-2; [21, 22]), the Developmental, Dimensional and Diagnostic Interview (3di; [24]) and the Diagnostic Interview for Social and Communication Disorders - 11th edition (DISCO-11; [25]). The Autism Diagnostic Interview - Revised (ADI-R) was not included, because, to our knowledge, no DSM-5 adapted algorithm for children and adolescents has been published, except for the adapted algorithm for children aged between 12 and 47 months [26]. Our decision to only include instruments with a wide age range was based on the desire to be as inclusive as possible in representing how ASD characteristics across a broad developmental span, while at the same time enabling comparisons across instruments.

Item mapping. Items included in the DSM-5 algorithms of the instruments (for ADOS-2, see [21, 22]); for 3di, see [27]; for DISCO, see [17]) were compared to the DSM-5 description of (sub-)criteria and exemplars, taking into account full item descriptions and coding options, and independent of the classification according to the instrument. Two raters (KE and JM), experienced in the diagnostic assessment of children with ASD, independently categorized all items. A multidisciplinary expert panel, consisting of KE, JM, IN, WM and AD, a group of professionals that is highly experienced in the diagnostic assessment of ASD both in the context of research and in clinical practice, discussed items when: (1) there was disagreement between the two raters, or (2) the categorization by the two raters was different from the categorization according to the instrument. Final decisions were based on the panel discussion. All expert panel members were trained in the assessment and coding of at least two of the three instruments.

Evaluation of algorithm classifications. Algorithm classification procedures of the evaluated instruments were compared to DSM-5 (sub-)criteria and diagnostic decision-making rules.

Identifying difficulties when operationalizing DSM-5 (sub-)criteria into behaviors. The discussion of the difficulties relating to the clarity of the DSM-5 criteria (as per aim 2, for results, see Discussion) was based on expert panel discussion. Items were discussed by the expert panel when there was disagreement between two raters, or when the categorization by two raters was different from the categorization according to the instrument. In addition, individual expert panel members also noted items for which they were unsure of the classification, and those items also contributed to the discussion.

Instruments

Autism Diagnostic Observation Schedule (ADOS-2). The ADOS-2 [21, 22] is a semi-structured, standardized observational assessment in which toys, activities and/or conversations are used to elicit communication, social interaction, play, and repetitive and stereotyped behaviors relevant to the diagnosis of ASD. Administration consists of direct observation by a trained examiner in a one-on-one situation (except for young children in the Toddler Module, and Modules 1 and 2, when a familiar adult is present as well). The ADOS-2 can be used to assess individuals from all ages and levels of functioning and offers five different modules and eight different algorithms, from which one module and algorithm is selected based on the individual's expressive language level and chronological age. Observations of the individual's behavior are coded on 28 to 41 items (depending on the module chosen), usually on a scale from 0 to 2, with higher scores indicating greater symptom severity. The administration of the ADOS-2 takes approximately 40 to 60 minutes. Revised DSM-5-adapted

algorithms were published in the ADOS-2 manual (for all modules apart from Module 4), with a sensitivity between .60 and .95, and specificity between .75 and 1 (depending on the administered module; [21, 22]). For Module 4 (for fluently verbal adolescents and adults), the DSM-5 algorithm is not integrated in the instruments' manual yet, but a research version has been published and demonstrated overall sensitivity between 84.6 and 90.5, and specificity between 72.1 and 82.2 [19, 20].

Developmental, Dimensional and Diagnostic Interview (3di). The 3di [24] is a computerized parental (or caregiver) interview that is a hybrid of a fully structured and a semi-structured interview. A trained examiner collects information of the individual's developmental history and of a broad range of skills and behaviors that are relevant for an ASD diagnosis, but also for co-occurring problems. Prior to the interview, the examiner imputes identifying information, which tailors the wording of questions. Scoring broader, complex questions is not required: Such questions were broken down into more specific items, to increase reliability [24]. The 3di was primarily designed to assess individuals aged 2-21 years with normal-range intellectual abilities, but it may also be used among those with intellectual disability and recently, an adult version of the interview was also published [28]. The 3di comprises more than 700 questions that are grouped in 23 different sections. The number of questions included in the interview has increased over the years, with different research groups adding new questions on specific DSM-5 related topics, hereby generating different parallel versions of the full interview instrument. Interviewers almost never administer every question: the 3di is constructed of different modules, each including a subset of questions. Depending upon the purpose and/or suspicion of co-occurring problems, the full autism module might be complemented with one of the modules on co-occurring problems. The majority of questions concerning atypical behaviors are coded on a 3-point severity scale: 0 (described behavior is not present), 1 (minimal evidence of described behavior), and 2 (definite or persistent evidence of described behavior). The 3di assumes the interviewer to rate whether behavior "ever" or "now" occurred. Administration time strongly depends on the selected module, ranging from 45 minutes (short version; [29]) to two hours. Psychometric properties for the DSM-5 version of 3di [27] have not yet been investigated. Classifications based on a preliminary version of this algorithm were compared to ADOS-2-classifications, showing a sensitivity of .84 and a specificity of .54 [18].

Diagnostic Interview for Social and Communication Disorders (DISCO-11). The DISCO-11 [25] is a semi-structured parental (or caregiver) interview, in which a trained examiner collects information about an individual's developmental history and a broad range of skills and behaviors relevant for an ASD diagnosis, but information on other domains is also collected. Individuals from all ages and levels of functioning can be assessed using the DISCO-11. The DISCO-11 comprises more than 300 items that are grouped in eight different sections.

The majority of items concerning atypical behaviors are coded on a 3-point severity scale: 0 (marked problem), 1 (minor problem), and 2 (no problem). For most of these items the DISCO-11 distinguishes both 'ever' and 'current' ratings of the individual's behavior. Other items are measuring the current level of functioning: the higher the level of achievement, the higher the score, with codes ranging between 0 and 12. Another type of items are about developmental milestones: for some, the actual age of achieving (in months) is coded, for others whether there was a delay in achieving specific developmental milestones. The last type of items rate the quality of behavior based on qualitative descriptions for each category (maximum of 10). Administering the complete interview takes approximately two to three hours, but it is also possible to only complete the items relevant for the diagnostic algorithms, resulting in a shortened administration time (about 45-60 min; [16]). The DSM-5 algorithm has been shown to have a good sensitivity and specificity, ranging from .85 to 1.00 and .74 to .89 respectively, based on different samples [17].

Results

Given that none of instruments' algorithms explicitly included items related to criterion D ('Significant impairment in functioning') or E ('Not better explained by intellectual disability'), our analyses focus on criterion A ('Deficits in social interaction and social communication'), B ('RRBIs'), and C ('Early onset').

Item mapping

All items were mapped onto each of the DSM-5 sub-criteria for ASD by the two coders (see Table 1 and Appendix B for the more detailed item mappings for each instrument). Inter-rater agreement was high, with agreement between expert raters for 68 out of 70 ADOS-2 items (97%), 62 out of 63 items for 3di (98%), and 80 of 85 DISCO-11 items (94%). An additional 23 items were discussed in the expert panel, as they were categorized differently by the raters compared to the instrument: for ADOS-2, 1 out of 70 items (1%; but note that the instrument only categorizes items based on the two main criteria and not based on sub-criteria); for 3di, 10 out of 63 items (16%); and for DISCO-11, 12 out of 85 items (14%).

[INSERT TABLE 1 HERE]

Autism Diagnostic Observation Schedule (ADOS-2). The ADOS-2 groups items into 'Social Affect' and 'Restrictive and Repetitive behaviors', without further specifications regarding DSM-5 sub-criteria. This division into 'Social Affect' and 'Restrictive and Repetitive behaviors' parallels our item mapping on criteria A and B, although one item was categorized differently: whereas the ADOS-2 manual categorized 'Reporting of

events' (in which the ability is evaluated to describe a non-routine event in an understandable manner, an item that is only included in Modules 3 and 4) under 'Social Affect', our item mapping did not organize this item in any of the sub-criteria, as it mainly reflects the level of expressive language skills, an aspect that is no longer part of the DSM-5 criteria.

Our analysis shows that the ADOS-2 DSM-5 algorithm mainly focuses on criterion A symptoms, and more specifically on A1 ('Deficits in socio-emotional reciprocity') and A2 ('Deficits in nonverbal communication'). Only 3 to 5 (out of 14 to 15) items cover the criterion B symptoms, with an emphasis on 'Stereotyped and repetitive behaviors' (B1). There are at most two ADOS-2 items measuring symptoms in the area of 'Deficits in relationships' (A3) and no items on 'Insistence on sameness and routines' (B2). Items both for A3 ('Deficits in relationships', for example, item 'Insight into social relationships') and for B2 (for example, item 'Compulsions or rituals') are available in the instruments, but these items are not included in most modules' algorithms. No indicators for early onset (criterion C) are available and the ADOS-2 focuses on current behaviors and does not include the presence of criterion A ('Deficits in social communication and interaction') or B ('Restricted, repetitive behaviors, interests or activities') symptoms in the past.

Developmental, Dimensional and Diagnostic Interview (3di). The developers of the 3di used clinical agreement to classify DSM-5 algorithm items into the specific sub-criteria, but factor analysis was not used for confirming this selection. We categorized 10 out of 63 items differently compared to the instrument. More details can be found in Appendix B.

All sub-criteria within A ('Deficits in social communication and interaction') and B ('Restricted, repetitive behaviors, interests or activities') are measured by at least five or more questions of the 3di. Multiple questions are used to assess the same exemplars. For example, three questions are included about sharing objects or food (A1, 'Deficits in social-emotional reciprocity'), five questions about stereotyped and repetitive speech (B1, 'Stereotyped and repetitive behaviors'), and seven questions about hypersensitivity to sounds (B4, 'Hyper- or hyporeactivity'). On the other hand, some exemplars are not covered, such as 'Failure to initiate or respond to social interactions' (under A1, 'Deficits in social-emotional reciprocity'), 'Difficulties with transitions' (under B2, 'Insistence on sameness and routines'), or 'Apparent indifference to pain/temperature' (under B4, 'Hyper- or hyporeactivity'). Even though the instrument comprises an extensive developmental history, no items on criterion C ('Early onset') are included in the algorithm. For all items, interviewers should take into account both current and past behaviors when attributing a score, matching the specification in DSM-5 that criteria can be met currently or by history, as long as the total presentation is currently impairing.

Diagnostic Interview for Social and Communication Disorders (DISCO-11). Based on clinical agreement, the developers of the algorithm mapped items from the DISCO to the DSM-5 sub-criteria based on clinical agreement and the item selection has not yet been validated using factor analyses [17]. Twelve out of 85 items were categorized differently in the current analyses compared to the original organization of items by the authors (for more details, see Appendix B).

All sub-criteria under criteria A and B are covered by seven or more DISCO items each. Some items are not applicable for younger children (<4y or <6/7y), these items are mainly related to 'Deficits in relationships' (A3). However, for younger children, six items remain applicable to measure sub-criterion A3. Within B1 ('Stereotyped and repetitive behaviors') and B2 ('Insistence on sameness and routines'), five and three items, respectively, cannot be coded for minimally verbal individuals, but all other items remain applicable. The different items belonging to a specific sub-criterion cover the full range of exemplars. However, more items are available in the interview that could be used to extend and maybe even improve the algorithms, in particular for younger and minimally verbal individuals. The algorithm also includes items on early onset (criterion C), and for most criterion A and B symptoms separate scoring of current and past behaviors is required.

Analysis of the algorithms

The development (i.e., the item selection procedure, the procedure used to set cut-off scores), decision-making rules and classification by the algorithm of the three instruments were compared to DSM-5 decision-making rules for ASD (see Table 2 and Table 3). In order to review the instruments' algorithms, the original authors' categorization of items under the DSM-5 sub-criteria was used. Therefore, the number of items for each sub-criterion shown in Table 2 might differ from Table 1 (mapping by our expert panel).

[INSERT TABLE 2 HERE]

[INSERT TABLE 3 HERE]

Autism Diagnostic Observation Schedule (ADOS-2). The DSM-5 algorithm of the ADOS-2 has been constructed by subdividing the standardization sample into five different groups based on age and verbal level corresponding with the five new ADOS-2 algorithms. Items were included based on their ability to distinguish individuals with autism from those without autism and comparability of concepts between modules [21, 22]. Algorithm items were subdivided in two different domains, 'Social Affect' (SA) and 'Restricted and Repetitive Behaviors' (RRB), based on exploratory and confirmatory factor analyses [30, 31].

To compute an ADOS-2 classification, all (recoded) A ('Deficits in social communication and interaction') and B ('Restricted, repetitive behavior, interests and activities') algorithm items are added and compared to one cut-off value. Such a classification procedure is not consistent with DSM-5 criteria and decision-making rules, as the ADOS-2 algorithm has no separate cut-off for criterion A (or SA) and criterion B (or RRB). An ADOS-2 classification of ASD can hence be provided based on criterion A ('Deficits in social communication and interaction') symptoms only. Furthermore, the skewed distribution of items over the different (sub-)criteria (see Table 2) could influence the final classification; for instance individuals with more severe problems in social-emotional reciprocity and nonverbal communication (and no RRBIs) are more likely to reach the threshold than individuals with less pronounced socio-communicative problems and many RRBIs.

ADOS-2 provides different cut-off scores for each module, and thus for different age groups and levels of ability. The instrument also implements some indices of severity. First, ADOS-2 distinguishes between overall cut-off scores for the classifications 'autism spectrum' and for 'autism', the latter referring to a more stringent cut-off. Second, overall raw total scores can be converted into a comparison score to estimate ASD symptom severity on a 10-point scale [21, 22]. Severity scores for domain totals (SA and RRB) are available in academic publications [32, 33], but are not included in the instrument's manual and might therefore be unknown to clinicians. Moreover, it is not yet clear how these specific severity scores for SA and RRB relate to the three severity levels for criteria A and B as described in DSM-5 [33, 34].

Developmental, Dimensional and Diagnostic Interview (3di). Different sets of items have been put forward to be included in the DSM-5 algorithms of the 3di [18, 27], but only one of those DSM-5 algorithms has been used in a peer-reviewed publication [27]. The full description of the algorithm was not included in the publication and is not integrated in the clinical software yet. Therefore, given the lack of transparency on how the algorithms were constructed and the underlying decision-making rules, it was necessary to obtain the algorithm from the authors of the instrument directly to carry out any analysis of it.

For the DSM-5 algorithm [27], 63 items were selected from the full set of items included in DSM-IV-TR algorithm [24], complemented by items from the Children's Communication Checklist (CCC, items that are also included in the full version of the 3di; [35]) via a two-stage process (see Table 2 and 3). In a first step, 3di subscales and items (belonging to the DSM-IV-TR algorithm plus items from the CCC, see Table C1) were selected based on their relevance with regards to DSM-5 behavior descriptions by the senior authors and developers of the original algorithm [27]. All (recoded) items were organized in a set of subscales and then organized under the DSM-5 sub-criteria (for the exact number of subscales and their names, see Table C1). In a second step, three items were

selected for each subscale, to reach the highest possible internal consistency, based on Cronbach's alpha. The same algorithm can be used for individuals of all levels of intellectual functioning under 18 years.

The 3di algorithm follows most of the DSM-5 decision-making rules for an ASD classification. Cut-offs were not based on statistical analyses, but based on consensus among authors [27]. First, cut-offs were determined for all subscales. Second, the threshold for the sub-criteria (A1, … , B4) was set on meeting the cut-off for at least one of the underlying subscales. Given the uneven distribution of subscales (and items) across the different sub-criteria, this decision-rule may have an effect on the classification; for example, the threshold for sub-criterion A1 ('Deficits in social-emotional reciprocity', with five subscales) is lower than the threshold for sub-criterion A3 ('Deficits in relationships', with three subscales) or B1 ('Stereotyped and repetitive behaviors', with two subscales). Third, and in line with DSM-5, a final classification of ASD requires scoring above the cut-off on all three sub-criteria of criterion A ('Deficits in social communication and interactions'), and two out of four sub-criteria within criterion B ('Restricted and repetitive behavior, interests, and activities'). Even though the 3di includes elaborate information on developmental history, no information on 'Presence of behaviors in early development' (criterion C) is included in the algorithm, and although both present and past presence of symptoms should be taken into account when rating, no explicit distinction is made between them. The 3di DSM-5 algorithm does not offer information on ASD severity.

Diagnostic Interview for Social and Communication Disorders (DISCO-11). For the DSM-5 algorithm of the DISCO-11 [17], 85 items were selected based on their relevance with regards to DSM-5 sub-criteria and exemplars (see Table 2 and 3). Item selection was done by researchers and reviewed by a panel of independent clinicians. The algorithm recodes original item codings into present or not present. Algorithm thresholds for sub-criteria were defined based on ROC curve analyses. This DSM-5 algorithm had comparable sensitivity and specificity across the different age and ability levels tested. The DISCO-11 strictly follows all DSM-5 decision-making rules for ASD classification. First, separate cut-offs for criteria A ('Deficits in social communication and interaction'), B ('Restricted and repetitive behavior, interests, and activities'), and C ('Early onset') are used. Second, all A sub-criteria and two out of four B sub-criteria have to be met to obtain a classification. As defined in DSM-5, behaviors based on current descriptions or by history are taken into account in the 'ever' classification of DISCO-11. A final ASD classification is only possible when all three criteria (A, B, and C) are met separately. Note that our item mapping suggested that most of the items included in criterion C are not consistent with DSM-5 (items related to development of language and pretend play, for more details see Item

mapping and Appendix D), which might have an effect on the classification. The DISCO DSM-5 algorithm does not offer any information on severity of ASD symptoms.

Discussion

The first aim of this study was to establish the content validity of three diagnostic assessment instruments in relation to the DSM-5 algorithms for ASD, namely ADOS-2 [21, 22], 3di [27], and DISCO-11 [17] and the second aim was to identify potential problems with the operationalization of DSM-5 diagnostic (sub-)criteria for ASD. Our analyses showed that the three instruments do not cover all ASD symptoms to the same extent and that their diagnostic classification procedures are not always in line with the DSM-5 ASD criteria. Furthermore, the interpretation of the DSM-5 behavioral A ('Deficits in social communication and interactions') and B ('Restricted and repetitive behavior, interests, and activities') criteria is sometimes ambiguous and the other criteria (C – 'Early onset', D – 'Significant impact on daily life functioning', and E – 'Not better explained by other developmental diagnosis') are not clearly defined.

The three instruments do not cover all ASD symptoms to the same extent

Differences in the nature of the instruments, their history and the development of the DSM-5 adapted algorithms can explain some of the variability in the symptoms included in the instruments' algorithms. More specifically, an observation scale such as ADOS-2 cannot include items on developmental history, whereas these items are available in the parental interviews, but not always included in the algorithms. In addition, the likelihood of observing less frequent, yet highly salient and clinically significant RRBIs is limited during the 45-minute time-window of the ADOS-2, which probably explains the under-representation of criterion B items in ADOS-2, compared to the interview instruments [33]. The absence of RRBIs in such a context should be interpreted with caution, as those behaviors might only occur under highly specific circumstances [33]. Similarly, authors state that it might be hard to capture deficits in building and maintaining relationships in a time-limited standardized observation, which could explain why this sub-criterion is (almost) absent in some ADOS-2 modules [33]. However, observation instruments provide the clinician with unique first-hand observations of the child. Whereas both parental interviews do a good job of representing all different sub-criteria, the ADOS-2 does not cover all sub-criteria (a finding that is in line with a previous item mapping by Huerta and colleagues [14]), which is partially due to the limitations of a time-constraint observation instrument.

The number of items included in the algorithms differed significantly across the three instruments. However, comparing the absolute number of items does not do justice to the instruments, as their items vary greatly in how broadly they are formulated. Whereas the items in ADOS-2 mostly refer to a broader area of functioning (e.g., 'Using gestures'), the 3di consists of highly specific questions (e.g., 'Shaking head for no'; 'Nodding head for yes'). In this regard, the DISCO-11 takes an intermediate position (e.g., 'Shaking or nodding head'). It is hence evident that ADOS-2 consists of fewer items than 3di or DISCO-11. Furthermore, compared to DISCO-11, the 3di emphasizes specific exemplars, which is probably partially related to the development of its DSM-5 algorithm, starting from existing subscales (see Results). Consequently, the 3di provides an elaborate picture of some specific DSM-5 exemplars, while other symptoms remain unexplored (e.g., seven out of ten B4 items focus on auditory sensitivity, but no items related to indifference to pain or temperature are included). Although exemplars do not represent an exhaustive list of symptoms within a specific criterion, the distribution of items across different exemplars is important to capture a range of different symptoms.

The nature and history of the different instruments can partly explain why not all DSM-5 criteria are represented in the different instruments. More items might also be required to capture the range of impairments in some sub-criteria. Although some empirical findings indeed suggest that it might be harder to capture socio-communicative problems in a few items compared to problems related to RRBIs [36], our item mapping demonstrated that instrument-specific mechanisms also play an important role: Whereas the algorithms of the ADOS-2 (as acknowledged by the authors of ADOS-2 in the manual) and the 3di consisted of more socio-communicative items compared to RRBI symptoms, the opposite pattern was found for the DISCO-11, where somewhat more items measured RRBI symptoms than problems with social interaction and communication. Moreover, as individual items may be more characteristic of particular subgroups of individuals, including a broader range of items could therefore improve sensitivity for different subgroups of individuals.

The differences in how the DSM-5 criteria are represented in the different instruments, and particularly the different limitations and advantages of parental interviews and observation scales, highlight that the combination of different diagnostic instruments increases their predictive value [6, 14]. Indeed, neither observation nor parental interviews should be the sole instrument used in diagnostic decision-making. At a minimum, clinicians should be aware of the limitations of specific instruments and use additional sources of information to address these limitations. For example, peer interactions are not evaluated by the ADOS-2 to provide an insight on peer relations. In this case, ADOS-2 information could be complemented with information from the semi-structured interviews.

The interpretation of DSM-5 behavioral criteria for ASD is sometimes ambiguous

The expert panel experienced some difficulties in assigning the items to the different (sub-)criteria, and the areas of greatest disagreement and discussion between raters are highlighted in this section. Taken together, our item mapping raised questions concerning the exact meaning of ASD symptoms as described in DSM-5, and their operationalization into concrete and measurable/observable behaviors. Within criterion A, the distinction between A1 ('Deficits in social-emotional reciprocity') and A3 ('Deficits in developing, maintaining and understanding relationships') appeared especially difficult [also see 14], as nearly all behaviors under A1 seem to be requirements for building and maintaining friendships (A3), although other reasons for deficits in A3 are possible as well. There does not only appear to be a hierarchical relationship between A1 and A3 symptoms, but they are also quite hard to distinguish from each other, as was reflected in the number of disagreements between our item mapping and the original placement of items for these sub-criteria. For instance, solely based on the sub-criteria and exemplars it is difficult to differentiate A1 exemplars 'not being able to maintain a reciprocal back-and-forth conversation' and 'a failure in the initiation or response to social behaviors' from A3 exemplars 'difficulties in adjusting behavior to suit various social contexts' and 'an absence of interest in peers'. Although some differentiation seems to be possible, individual items could equally map across these different descriptions and therefore map to A1 and A3. Hence it might be difficult to unravel and separately measure these different symptoms in research.

A considerable proportion of individuals with ASD is nonverbal or minimally verbal [37, 38]. However, the distinction between A2 and A1 can be especially difficult in this group. For example, when a nonverbal individual does not point to share an interest, should that behavior be considered as a problem in nonverbal communicative behavior used for social interaction (A2) or as a deficit in social-emotional reciprocity (A1)? Behaviors like joint attention or sharing enjoyment are all nonverbal social behaviors. Clear guidelines are lacking on how to differentiate A1 and A2 in a population that uses nonverbal behaviors as their primary mode of communication. A1 focuses on reciprocity, regardless of the modality (verbal or nonverbal), whereas A2 covers the quantity and quality of nonverbal behaviors serving the social interaction. Diagnostic instruments should try to distinguish those aspects of (non)verbal behaviors in different items or questions.

Moreover, it appeared sometimes problematic to distinguish between symptoms related to B1 ('Stereotyped and repetitive behaviors'), B2 ('Insistence on sameness') and B3 ('Highly restricted, fixated interests'). For example, a child with an especially strong interest in a specific animation series (B3), who imitates

entire conversations from that series (B1) and insists on watching the series every evening at seven o'clock (B2), could reach threshold on three B criteria based on one fixated interest that prevails in other aspects of functioning. In these cases, it remains unclear how clinicians or researchers should categorize such complex behaviors. On the one hand, it appears unfair to code one set of behaviors under multiple sub-criteria, as individuals will reach diagnostic thresholds – if impairing across contexts – very quickly based on one complex behavior. On the other hand, guidelines are lacking on which sub-criteria should be prioritized over others in these instances.

Taken together, mapping the instruments' items onto DSM-5 diagnostic criteria revealed difficulties in the operationalization in clear, measurable or observable behaviors and the distinction between specific sub-criteria. By no means are we pleading for a checklist of concrete symptoms that have to be met. Diagnostic ASD evaluations should comprise an extensive assessment of the individual in various contexts, comparing the individual's behaviors not only to the diagnostic criteria as described in manuals, but also taking into account expectations based on the overall level of intellectual functioning [2]. However, it appears important to reformulate and clarify some of the symptoms enlisted in DSM-5, such that researchers and clinicians can reach consensus about how to clearly map behaviors that are part of the ASD phenotype onto the DSM-5 sub-criteria.

Different classification procedures in the DSM-5 algorithms can lead to different algorithm outcomes based on the three instruments

The development of the algorithms differs between the instruments, both with regards to the selection of items, and with regards to determining the cut-off. The DISCO-11 selected items in a fully top-down manner, including items with the highest content validity with regards to DSM-5 criteria. The ADOS-2 and 3di took another approach, integrating bottom-up (data-driven) and top-down (construct-driven) elements. Instruments also significantly differed with regards to how the cut-off was set. For ADOS-2 and DISCO-11, cut-off scores were determined based on ROC analyses, whereas for 3di it was based on consensus in the research team.

Psychometric properties of ADOS-2 and DISCO-11 have been reported to be good to very good [17, 21, 22], while the sensitivity and specificity of the DSM-5 algorithm of the 3di have not yet been established. To date, psychometric research for the three instruments has been based on groups with a DSM-IV-TR diagnosis of autism. Research including individuals with a clinical DSM-5 diagnoses of ASD is needed to establish psychometric properties, but this is largely lacking to our knowledge (except for two studies showing good psychometric properties of ADOS-2 in adults; [39, 40]).

For criterion A ('Social interaction and communication') and B ('RRBIs'), the classification procedures of 3di and DISCO-11 are in line with the decision-rules as described in DSM-5, in the sense that an ASD classification requires combined problems in criterion A and B. However, criterion C ('Early onset') is less well implemented in the parental interviews, and interviewers need to remain mindful about this criterion. Although the 3di contains developmental items, its DSM-5 algorithm does not include any items about developmental history. DISCO-11 includes a set of items of which several are not part of DSM-5, which could have an impact on diagnostic classification according to the instrument.

Based on the ADOS-2 algorithm, however, an ASD classification can be given without meeting all DSM-5 criteria. For example, individuals with social communication disorder who only show impairments in criterion A ('Deficits in social communication and interaction'), but no RBBIs (criterion B), may also reach the threshold for ASD classification on ADOS-2. The ADOS-2 manual should stipulate more clearly the potential consequence of this classification rule on diagnostic outcome according to the instrument. In both academic and clinical research, especially in the USA, the ADOS-2 is widely used, typically in combination with Autism Diagnostic Interview-Revised (ADI-R; [41]), and they are often referred to as the *gold standard* for ASD diagnosis. However, the psychometric properties of the (combined) ADOS-2 and ADI-R for the clinical diagnosis of ASD in the DSM-5 era, has not been studied yet.

The interpretation of DSM-5 criteria C ('Early onset'), D ('Impact on daily life functioning'), and E ('Not better explained by intellectual disability'), and the three levels of severity has not yet been clearly defined

Our comparative analyses also revealed that criteria C, D, and E were not (sufficiently) included in the commonly-used diagnostic instruments. These criteria seem to be neglected and underspecified in the development of instruments and published work at this stage. Compared to DSM-IV-TR, the early onset criterion is less strictly defined and the presence of symptoms that may be masked in early development, but cause impairments later-on, is clearly acknowledged in criterion C. However, our analyses showed that diagnostic instruments did not integrate this less strictly defined age-of-onset criterion in their algorithms. Clinicians generally use other sources of information to establish criterion C. Although the impairing effect of ASD symptoms on important areas of functioning (criterion D) is not explicitly included in the instruments' algorithms, all instruments provide some information on the impact of characteristics that should be interpreted by the clinician, taking into account all available information. However, estimating the impairing impact can be rather difficult, certainly in individuals with more subtle, or with co-occurring problems. In addition, DSM-5 stipulates that disturbances should not be

better explained by intellectual disability or global developmental delay (criterion E), but no guidance is provided on how to make this determination.

DSM-5 explicitly refers to the heterogeneous environmental modifications that are required for daily functioning, by outlining three levels of severity for socio-communicative impairments and RRBIs separately, based on the amount of support needed. Qualitative descriptions of the different levels of support are provided in DSM-5, but the operationalization seems to be more closely related to severity of ASD symptoms. However, previous research suggests that there is little overlap between different concepts related to severity and support [42]. Without a clear conceptualization and a standard method it hence seems unlikely that professionals are consistent in their classification as requiring 'support', 'substantial support' or 'very substantial support'. Furthermore, there are no guidelines on whether characteristics with a known impact on severity or level of support, such as age, cognitive level, language ability or adaptive behavior, should be taken into account, nor on how this should be done [34, 42].

Diagnostic ASD assessment instruments and DSM-5: Conclusions and recommendations for research and clinical practice

Based on growing empirical evidence, DSM-5 has abandoned the different sub-classifications within the autism spectrum, and has stipulated two (instead of three) core domains of impairment [18, 30, 31, 43–45]. Inter-individual variability has been incorporated – at least partially – by including a range of exemplars and the option to indicate the level of required support separately for symptoms related to socio-communication problems and RRBIs. Furthermore, the co-occurrence of ASD with other disorders has now been acknowledged. In all, the new DSM has improved greatly with respect to clarifying nosology and providing more transparent descriptions of the core ASD characteristics [46, 47].

Diagnostic assessment instruments that have developed specific DSM-5 algorithms differ greatly from each other with respect to which ASD features are measured and their compatibility with the DSM-5 classification rules. It is crucial that users understand these limitations, both in terms of ASD characteristics (not) covered and in terms of the classification according to the algorithm, given the importance of these instruments in the context of academic research and clinical diagnostic assessments. Clinicians using the instruments in the context of diagnostic decision-making take into account that – as before – diagnostic classification should never be based solely on the score on one (or two) instruments [also see, e.g., 6], but should rely on the integration and clinical interpretation of different sources of information by a multidisciplinary team of experienced clinicians [2, 5]. Our

results demonstrate which sub-criteria are (not) covered by specific diagnostic instruments, and hence highlight areas for each instrument where it would be important to collect additional information. In the context of (international) research, it should be emphasized that DSM-5 criteria are implemented differently by ADOS-2, 3di and DISCO-11. Caution is therefore warranted, when using one or two instruments to *validate* a clinical diagnosis, as the exclusion of participants based on the outcome of these instruments might lead to a biased understanding.

In order to advance clinical practice and research we recommend future work be directed towards solving some of the existing ambiguities with regards to the definition and measurement of impairment in current functioning (criterion D) and severity levels. In addition, it remains rather difficult to distinguish between sub-criteria within criterion A and B, especially in the context of very young children or nonverbal individuals. Providing more clarity could lead to a more stable and accurate classification. Note that to fully understand and situate the symptoms, it is important that professionals carefully read the full text supplementing the list of criteria.

DSM-5 is currently the most important classification manual in the context of ASD. Recently, the International Classification of Diseases (ICD) has published its novel guidelines [13] and the criteria show a strong parallel with DSM-5. However, there also are notable differences between both manuals. In contrast to DSM-5, ICD-11 enlists eight subcategories in ASD, based on co-occurring intellectual or language impairments. In addition, ICD-11 does not provide concrete exemplars, and no required combination of number of symptoms. On the one hand, this might give more flexibility to clinicians, who have to judge whether an individual meets the threshold. On the other hand, this could also negatively impact the reliability and stability of diagnoses across settings and professionals. The future will tell how the differences between both classification systems will be integrated in diagnostic assessment instruments, and how they will impact prevalence rates.

Conflict of interest

Some of the authors (SL, SC, JM, IN) contributed to the development of the DSM-5 algorithm of the DISCO-11, and some of them are also involved in the DISCO-11 training for researchers (SL, SC, JM, IN, KE).

References

1. Ashwood KL, Buitelaar J, Murphy D, Spooren W, Charman T (2015) European clinical network: autism spectrum disorder assessments and patient characterisation. Eur Child Adolesc Psychiatry 24:985–995 . https://doi.org/10.1007/s00787-014-0648-2

2. Constantino JN, Charman T (2016) Diagnosis of autism spectrum disorder: Reconciling the syndrome, its diverse origins, and variation in expression. Lancet Neurol 15:279–291 . https://doi.org/10.1016/S1474-4422(15)00151-9

3. Falkmer T, Anderson K, Falkmer M, Horlin C (2013) Diagnostic procedures in autism spectrum disorders: a systematic literature review. Eur Child Adolesc Psychiatry 22:329–340 . https://doi.org/10.1007/s00787-013-0375-0

4. Gardner L, Erkfritz-Gay K, Campbell JM, Bradley T, Murphy L (2016) Purposes of Assessment. In: Matson JL (ed) Handbook of Assessment and Diagnosis of Autism Spectrum Disorder. Springer International Publishing, Cham, pp 27–43

5. National Institute for Health and Care Excellence (2011) Autism spectrum disorder in under 19s: recognition, referral and diagnosis (NICE guideline CG128). http:// nice.org.uk/guidance/cg128

6. Kim SH, Lord C (2012) Combining information from multiple sources for the diagnosis of autism spectrum disorders for toddlers and young preschoolers from 12 to 47 months of age. J Child Psychol Psychiatry 53:143–151 . https://doi.org/10.1111/j.1469-7610.2011.02458.x

7. American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders (5th edition). Author, Washington, D.C.

8. American Psychiatric Association (2000) Diagnostic and statistical manual of mental disorders (4th edition text revision). Author, Washington, D.C.

9. McPartland JC, Reichow B, Volkmar FR (2012) Sensitivity and specificity of proposed DSM-5 diagnostic criteria for autism spectrum disorder. J Am Acad Child Adolesc Psychiatry 51:368–383 . https://doi.org/10.1016/j.jaac.2012.01.007

10. Bennett M, Goodall E (2016) A Meta-analysis of DSM-5 Autism Diagnoses in Relation to DSM-IV and DSM-IV-TR. Rev J Autism Dev Disord 3:119–124 . https://doi.org/10.1007/s40489-016-0070-4

11. Kulage KM, Smaldone AM, Cohn EG (2014) How will DSM-5 affect autism diagnosis? A systematic literature review and meta-analysis. J Autism Dev Disord 44:1918–1932 . https://doi.org/10.1007/s10803-014-2065-2

12. Sturmey P, Dalfern S (2014) The Effects of DSM5 Autism Diagnostic Criteria on Number of Individuals Diagnosed with Autism Spectrum Disorders: A Systematic Review. Rev J Autism Dev Disord 1:249–252 . https://doi.org/10.1007/s40489-014-0016-7

13. World Health Organization (2018) International statistical classification of diseases and related health problems (11th Revision). https://icd.who.int/browse11/l-m/en

14. Huerta M, Bishop SL, Duncan A, Hus V, Lord C (2012) Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. Am J Psychiatry 169:1056–1064 . https://doi.org/10.1176/appi.ajp.2012.12020276

15. Barton ML, Robins DL, Jashar D, Brennan L, Fein D (2013) Sensitivity and Specificity of Proposed DSM-5 Criteria for Autism Spectrum Disorder in Toddlers. J Autism Dev Disord 43:1184–1195 . https://doi.org/10.1007/s10803-013-1817-8

16. Carrington S, Leekam S, Kent R, Maljaars J, Gould J, Wing L, Le Couteur A, Van Berckelaer-Onnes I, Noens I (2015) Signposting for diagnosis of Autism Spectrum Disorder using the Diagnostic Interview for Social and Communication Disorders (DISCO). Res Autism Spectr Disord 9:45–52 . https://doi.org/10.1016/j.rasd.2014.10.003

17. Kent R, Carrington S, Couteur A, Gould J, Wing L, Maljaars J, Noens I, Berckelaer-Onnes I, R Leekam S (2013) Diagnosing Autism Spectrum Disorder: who will get a DSM-5 diagnosis? J Child Psychol Psychiatry 54:1242–1250 . https://doi.org/10.1111/jcpp.12085

18. Slappendel G, Mandy W, Ende J van der, Verhulst FC, Sijde A van der, Duvekot J, Skuse D, Greaves-Lord K (2016) Utility of the 3Di Short Version for the Diagnostic Assessment of Autism Spectrum Disorder and Compatibility with DSM-5. J Autism Dev Disord 46:1834–1846 . https://doi.org/10.1007/s10803-016-2713-9

19. Hus V, Lord C (2014) The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm and Standardized Severity Scores. J Autism Dev Disord 44:1996–2012 . https://doi.org/10.1007/s10803-014-2080-3

20. Pugliese CE, Kenworthy L, Bal VH, Wallace GL, Yerys BE, Maddox BB, White SW, Popal H, Armour AC, Miller J, Herrington JD, Schultz RT, Martin A, Anthony LG (2015) Replication and Comparison of the Newly Proposed ADOS-2, Module 4 Algorithm in ASD Without ID: A Multi-site Study. J Autism Dev Disord 45:3919–3931 . https://doi.org/10.1007/s10803-015-2586-3

21. Lord C, Luyster R, Gotham K, Guthrie W (2012) Autism Diagnostic Observation Schedule, second edition (ADOS-2): Manual. Part II: Toddler Module. Western Psychological Services, Torrance, CA

22. Lord C, Luyster R, Gotham K, Guthrie W (2012) Autism Diagnostic Observation Schedule, second edition (ADOS-2): Manual. Part I: Modules 1-4. Western Psychological Services, Torrance, CA

23. Randall M, Egberts KJ, Samtani A, Scholten RJPM, Hooft L, Livingstone N, Sterling-Levis K, Woolfenden S (2018) Diagnostic tests for autism spectrum disorder (ASD) in preschool children. Cochrane Database Syst Rev 7: . https://doi.org/10.1002/14651858.CD009044.pub2

24. Skuse D, Warrington R, Bishop D, Chowdhury U, Lau J, Mandy W, Place M (2004) The developmental, dimensional and diagnostic interview (3di): a novel computerized assessment for autism spectrum disorders. J Am Acad Child Adolesc Psychiatry 43:548–558 . https://doi.org/10.1097/00004583-200405000-00008

25. Wing L (2006) Diagnostic Interview for Social and Communication Disorders, 11th edition (DISCO-11). Centre for Social and Communication Disorders, Bromley, United Kingdom

26. Kim SH, Thurm A, Shumway S, Lord C (2013) Multisite Study of New Autism Diagnostic Interview-Revised (ADI-R) Algorithms for Toddlers and Young Preschoolers. J Autism Dev Disord 43:1527–1538 . https://doi.org/10.1007/s10803-012-1696-4

27. Mandy W, Wang A, Lee I, Skuse D (2017) Evaluating social (pragmatic) communication disorder. J Child Psychol Psychiatry 58:1166–1175 . https://doi.org/10.1111/jcpp.12785

28. Mandy W, Clarke K, McKenner M, Strydom A, Crabtree J, Lai M-C, Allison C, Baron-Cohen S, Skuse D (2017) Assessing Autism in Adults: An Evaluation of the Developmental, Dimensional and Diagnostic Interview—Adult Version (3Di-Adult). J Autism Dev Disord 1–12 . https://doi.org/10.1007/s10803-017-3321-z

29. Santosh PJ, Mandy WPL, Puura K, Kaartinen M, Warrington R, Skuse DH (2009) The construction and validation of a short form of the developmental, diagnostic and dimensional interview. Eur Child Adolesc Psychiatry 18:521–524 . https://doi.org/10.1007/s00787-009-0004-0

30. Gotham K, Risi S, Dawson G, Tager-Flusberg H, Joseph R, Carter A, Hepburn S, McMahon W, Rodier P, Hyman SL, Sigman M, Rogers S, Landa R, Spence MA, Osann K, Flodman P, Volkmar F, Hollander E, Buxbaum J, Pickles A, Lord C (2008) A replication of the Autism Diagnostic Observation Schedule (ADOS) revised algorithms. J Am Acad Child Adolesc Psychiatry 47:642–651 . https://doi.org/10.1097/Chi.0b013e31816bffb7

31. Gotham K, Risi S, Pickles A, Lord C (2007) The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. J Autism Dev Disord 37:613–627 . https://doi.org/10.1007/s10803-006-0280-1

32. Esler AN, Bal VH, Guthrie W, Wetherby A, Weismer SE, Lord C (2015) The Autism Diagnostic Observation Schedule, Toddler Module: Standardized Severity Scores. J Autism Dev Disord 45:2704–2720 . https://doi.org/10.1007/s10803-015-2432-7

33. Hus V, Gotham K, Lord C (2014) Standardizing ADOS Domain Scores: Separating Severity of Social Affect and Restricted and Repetitive Behaviors. J Autism Dev Disord 44:2400–2412 . https://doi.org/10.1007/s10803-012-1719-1

34. Weitlauf AS, Gotham KO, Vehorn AC, Warren ZE (2014) Brief Report: DSM-5 "Levels of Support:" A Comment on Discrepant Conceptualizations of Severity in ASD. J Autism Dev Disord 44:471–476 . https://doi.org/10.1007/s10803-013-1882-z

35. Bishop DV (1998) Development of the Children's Communication Checklist (CCC): a method for assessing qualitative aspects of communicative impairment in children. J Child Psychol Psychiatry 39:879–891

36. Bishop SL, Havdahl KA, Huerta M, Lord C (2016) Sub-dimensions of social-communication impairment in autism spectrum disorder. J Child Psychol Psychiatry 57:909–916 . https://doi.org/10.1111/jcpp.12510

37. Norrelgen F, Fernell E, Eriksson M, Hedvall Å, Persson C, Sjölin M, Gillberg C, Kjellmer L (2015) Children with autism spectrum disorders who do not develop phrase speech in the preschool years. Autism Int J Res Pract 19:934–943 . https://doi.org/10.1177/1362361314556782

38. Tager-Flusberg H, Kasari C (2013) Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. Autism Res Off J Int Soc Autism Res 6:468–478 . https://doi.org/10.1002/aur.1329

39. Fusar-Poli L, Brondino N, Rocchetti M, Panisi C, Provenzani U, Damiani S, Politi P (2017) Diagnosing ASD in Adults Without ID: Accuracy of the ADOS-2 and the ADI-R. J Autism Dev Disord 47:3370–3379 . https://doi.org/10.1007/s10803-017-3258-2

40. Maddox BB, Brodkin ES, Calkins ME, Shea K, Mullan K, Hostager J, Mandell DS, Miller JS (2017) The Accuracy of the ADOS-2 in Identifying Autism among Adults with Complex Psychiatric Conditions. J Autism Dev Disord 47:2703–2709 . https://doi.org/10.1007/s10803-017-3188-z

41. Rutter M, Lord C (2003) Autism Diagnostic Interview, Revised. Western Psychological Services, Los Angeles, CA

42. Mehling MH, Tassé MJ (2016) Severity of Autism Spectrum Disorders: Current Conceptualization, and Transition to DSM-5. J Autism Dev Disord 46:2000–2016 . https://doi.org/10.1007/s10803-016-2731-7

43. Lord C, Petkova E, Hus V, Gan W, Lu F, Martin DM, Ousley O, Guy L, Bernier R, Gerdts J, Algermissen M, Whitaker A, Sutcliffe JS, Warren Z, Klin A, Saulnier C, Hanson E, Hundley R, Piggot J, Fombonne E, Steiman M, Miles J, Kanne SM, Goin-Kochel RP, Peters SU, Cook EH, Guter S, Tjernagel J, Green-Snyder LA, Bishop S, Esler A, Gotham K, Luyster R, Miller F, Olson J, Richler J, Risi S (2012) A multisite study of the clinical diagnosis of different autism spectrum disorders. Arch Gen Psychiatry 69:306–313 . https://doi.org/10.1001/archgenpsychiatry.2011.148

44. Mandy W, Charman T, Puura K, Skuse D (2014) Investigating the cross-cultural validity of DSM-5 autism spectrum disorder: evidence from Finnish and UK samples. Autism Int J Res Pract 18:45–54 . https://doi.org/10.1177/1362361313508026

45. Mandy WPL, Charman T, Skuse DH (2012) Testing the Construct Validity of Proposed Criteria for DSM-5 Autism Spectrum Disorder. J Am Acad Child Adolesc Psychiatry 51:41–50 . https://doi.org/10.1016/j.jaac.2011.10.013

46. Clarke DE, Narrow WE, Regier DA, Kuramoto SJ, Kupfer DJ, Kuhl EA, Greiner L, Kraemer HC (2013) DSM-5 field trials in the United States and Canada, Part I: study design, sampling strategy, implementation, and analytic approaches. Am J Psychiatry 170:43–58 . https://doi.org/10.1176/appi.ajp.2012.12070998

47. Lord C, Jones RM (2012) Annual research review: re-thinking the classification of autism spectrum disorders. J Child Psychol Psychiatry 53:490–509 . https://doi.org/10.1111/j.1469-7610.2012.02547.x

Table 1.

Summary of ADOS-2, 3di and DISCO-11 item mappings on DSM-5 (sub-)criteria for ASD [a]

DSM-5 (sub-)criteria	**ADOS-2** *n* of items [b]	**3di** *n* of items	**DISCO-11** *n* of items
A - Deficits in social communication and interaction	**9-11**	**35**	**36**
A1 - Deficits in social-emotional reciprocity	5-6	8	10
A2 - Deficits in nonverbal communication	3-5	12	14
A3 - Deficits in relationships	0-2	15	12[d]
B - Restricted, repetitive behavior, interests or activities	**3-5**[c]	**27**	**40**
B1 - Stereotyped / repetitive behaviors	2-3	6	13[e]
B2 - Insistence on sameness and routines	0	5	10[e]
B3 - Restricted, fixated interests	1	6	7
B4 - Hyper- or hyporeactivity to sensory input	1	10	10
C - Early onset	**n.a.**	**0**	**2**
Other items	**0-1**	**1**	**7**
Total *n* items	**14-15**	**63**	**85**

[a] Overview of concrete item mappings for all three instruments are available in Appendix B; [b] Ranges refer to the differences between the different Modules; [c] In some Modules one item (Repetitive behaviors or interests) was placed under two sub-criteria (B1 and B3); [d] Some items are not applicable for children <4y; [e] Some items are not applicable for minimally verbal individuals.

Table 2.

ADOS-2, 3di and DISCO-11 algorithm computation and classification compared to DSM-5 criteria and sub-criteria for ASD.

DSM-5 (sub)criteria	ADOS-2	3di	DISCO-11
A - Deficits in social communication and interaction	SA: no separate cut-off	A: 3/3 sub-criteria	A: 3/3 sub-criteria
A1 - Deficits in social-emotional reciprocity	-	1/5 subscales [b]	A1: 3/10 items
A2 - Deficits in nonverbal communication	-	1/4 subscales	A2: 1/14 items
A3 - Deficits in relationships	-	1/3 subscales	A3: 3/14 items
B - Restricted, repetitive behavior, interests or activities	RRB: no separate cut-off	B: 2/4 sub-criteria	B: 2/4 sub-criteria
B1 - Stereotyped / repetitive behaviors	-	B1: 1/2 subscales	B1: 1/13 items
B2 - Insistence on sameness and routines	-	B2: 1/2 subscales	B2: 1/11 items
B3 - Restricted, fixated interests	-	B3: 1/2 subscales	B3: 1/6 items
B4 - Hyper- or hyporeactivity to sensory input	-	B4: 1/2 subscales	B4: 1/10 items
C - Early onset	n.a.	n.a.	C: 1/7 items
ASD Classification	SA+RRB: meet overall cut-off	A+B: meet both criteria	A+B+C: meet all 3 criteria[a]
ASD Severity	CS: scale 1-10 (SA, RRB, and total)	-	**-**

SA = Social Affect; RRB = Restricted and Repetitive Behaviors; CS = Comparison Score

[a] Classification can be based on behaviors currently and by history ('ever' classification) or current behaviors only ('current' classification). [b] An overview of the different subscales, can be found in Appendix C.

Note. Whereas the ADOS-2 has an overall cut-off score and no separate cut-offs for the different sub-criteria, the 3di and the DISCO-11 have separate cut-offs for the different sub-criteria. The 3di algorithm consists of a hierarchical system, where different items (questions) are grouped into subscales. For every subscale, a separate cut-off is determined. Multiple subscales are grouped under one DSM-5 sub-criterion. To meet the cut-off of a DSM-5 sub-criterion, individuals have to meet cut-off of at least one of the subscales belonging to that sub-criterion. For the DISCO-11, individuals have to meet threshold on a set of items (questions) in order to score above cut-off on a specific sub-criterion

Table 3.

Summary of the comparison of the DSM-5 algorithms of ADOS-2, 3di and DISCO-11

	ADOS-2	3di	DISCO-11
Item-selection procedure?	Bottom-up: Statistical analyses determined which items provided the best distinction between individuals with versus without autism.	Top-down and bottom-up: A panel discussion selected the items that represented the DSM-5 criteria in the best way. Afterwards, three items were selected such that scales reached the highest internal consistency.	Top-down: A panel discussion selected the items that represented the DSM-5 criteria in the best way.
How was threshold determined?	Based on ROC curves	Based on clinical judgement	Based on ROC curves
Separate threshold for Criteria A and B?	No	Yes	Yes
Criterion C included in algorithm?	No	No	Yes
Severity score available?	Yes	No	No
Algorithm details published?	Yes	No	Yes
Same algorithm for all children?	No: different modules based on age and expressive language	Yes	Yes

Note. Criterion A = Deficits in social communication and interaction; Criterion B = Restricted, repetitive behavior, interests or activities; Criterion C = Early onset.

Appendix list

Appendix A. DSM-5 diagnostic criteria and sub-criteria for autism spectrum disorder

Table A1.

DSM-5 diagnostic criteria and sub-criteria for autism spectrum disorder 299.00 [1, pp. 50-51]

A	**Persistent deficits in social communication and social interaction across multiple contexts, as manifested by the following, currently or by history (examples are illustrative, not exhaustive; see text):**
A1	Deficits in social-emotional reciprocity
Exemplars	*Ranging from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions*
A2	Deficits in nonverbal communicative behaviors used for social interaction
Exemplars	*Ranging from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication*
A3	Deficits in developing, maintaining, and understanding relationships
Exemplars	*Ranging from difficulties adjusting behavior to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers*
B	**Restricted, repetitive patterns of behavior, interests, or activities, as manifested by at least two of the following, currently or by history (examples are illustrative, not exhaustive; see text):**
B1	Stereotyped or repetitive motor movements, use of objects, or speech
Exemplars	*e.g., simple motor stereotypes, lining up toys or flipping objects, echolalia, idiosyncratic phrases*
B2	Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behavior
Exemplars	*e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat same food every day*
B3	Highly restricted, fixated interests that are abnormal in intensity or focus
Exemplars	*e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interests*
B4	Hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of the environment
Exemplars	*e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement*
C	**Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life).**
D	**Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning.**
E	**These disturbances are not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay. Intellectual disability and autism spectrum disorder frequently co-occur; to make comorbid diagnoses of autism spectrum disorder and intellectual disability, social communication should be below that expected for general developmental level.**

Appendix B. Item mapping of the different diagnostic instruments

All items included in the DSM-5 algorithms of the three diagnostic assessment instruments were mapped onto each of the DSM-5 sub-criteria for ASD (see Table B1, B2, and B3, for the ADOS-2, 3di, and DISCO-11, respectively). The general description of our findings can be found in the main text of the manuscript. In this Appendix, a detailed discussion of the items that were categorized differently compared to the instrument follows each table.

Table B1.

Items included in the DSM-5 algorithm of ADOS-2 were mapped onto DSM-5 sub-criteria

DSM-5 criteria	**Item description**	**Modules**						
		T no words	**T** some words	**I** no words	**I** some words	**2**	**3**	**4**
A1	Frequency of vocalization	A2		A2	A2			
	Conversation						A8	A8
	Shared enjoyment	B6		B5	B5	B3	B4	
	Showing	B12		B9	B9	B5		
	Communicating own emotions							B5
	Response name		B7					
	Ignoring		B8					
	Requesting		B9					
	Initiation joint attention	B13	B13	B10	B10	B6		
	Response joint attention	B14		B11				
	Quality social overtures	B15	B15	B12	B12	B8	B7	B9
	Amount social overtures (parent)		B16b					
	Amount social overtures (general)							B10
	Quality social response						B9	
	Amount reciprocal social communication					B11	B10	B12
	***n* of items**	**6**	**6**	**6**	**5**	**5**	**5**	**5**
A2	Pointing		A7		A7	A6		
	Gestures	A8		A8	A8	A7	A9	
	Emotional gestures							A10
	Unusual eye contact	B1	B1	B1	B1	B1	B1	B1
	Facial expressions	B4	B4	B3	B3	B2	B2	B2
	Integration gaze other behaviors	B5	B5	B4	B4			
	***n* of items**	**4**	**4**	**4**	**5**	**4**	**3**	**3**
A3	Insight in relationships							B7
	Overall quality rapport		B18			B12	B11	B13
	***n* of items**	**0**	**1**	**0**	**0**	**1**	**1**	**2**
B1	Intonation	A3		A3				A2
	Stereotyped / idiosyncratic language				A5	A4	A4	A4
	Hand / finger / complex mannerisms	D2	D2	D2	D2	D2	D2	D2
	Repetitive interests / *behaviors*	D5	D5	D4	D4	D4		
	***n* of items**	**3**	**2**	**3**	**3**	**3**	**2**	**3**
B2								
	***n* of items**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
B3	Repetitive *interests* / behaviors	D5	D5	D4	D4	D4		
	Excessive interests						D4	D4
	***n* of items**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
B4	Unusual sensory interests	D1	D1	D1	D1	D1	D1	D1
	***n* of items**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
C		n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Other	Reporting of events						*A7*	
	n of items	**0**	**0**	**0**	**0**	**0**	**1**	**0**

Note. A1 = Deficits in social-emotional reciprocity; A2 = Deficits in nonverbal communication; A3 = Deficits in relationships; B1 = Stereotyped/repetitive behaviors; B2 = Insistence on sameness and routines; B3 = Restricted, fixated interests; B4 = Hyper- or hyporeactivity to sensory input; Criterion C = Early onset.

The ADOS-2 manual [2, 3] groups items into 'Social Affect' and 'Restrictive and Repetitive behaviors', without further specifications regarding DSM-5 sub-criteria. This division into 'Social Affect' and 'Restrictive and Repetitive behaviors' parallels our item mapping on Criteria A ('Deficits in social communication and interaction') and B ('Restricted, repetitive behavior, interests or activities'), although one item was categorized differently (see Table B1). Whereas the ADOS-2 manual categorized 'Reporting of events' (only included in Module 3) under 'Social Affect', our item mapping did not organize this item in any of the sub-criteria, as it mainly reflects the level of expressive language skills, an aspect that is no longer part of the DSM-5 criteria [1].

Table B2.

Items included in the DSM-5 algorithm of 3di were mapped onto DSM-5 sub-criteria

DSM-5 criterium	**Item description**	**Item number**
A1	Understanding of parental tone of voice	223
	Sharing food with peers	299
	Sharing possessions with peers	300
	Sharing without prompting	301
	Seeking comfort if emotionally upset	314
	Talks repetitively about things no-one is interested in	679
	Does not seem to realise need to explain what is talking about	688
	Ignores conversational cues	705
	***n* of items**	**8**
A2	Response to parental facial expression *(A1)**	224
	Getting eye contact across room	248
	Smiles in greeting after separation	251
	Smiles to familiar outside house	252
	Claps to indicate appreciation	279
	Nods agree	284
	Shakes head	285
	Reads facial expression of someone outside family	626
	Does not read facial expressions or tone of voice	707
	Does not use facial expression to convey feelings	708
	Avoids eye contact in conversation	709
	Looks away in conversation	710
	***n* of items**	**12**
A3	Regarded as rude *(A1)**	237
	Does not vary behaviour according to social context *(A1)**	317
	Less socially aware than peers *(A1)**	322
	Inappropriate social behavior outside home *(A1)**	323
	Engaged in pat-a-cake etc as infant	330
	Engages in joint play requiring cooperation	331
	Has best friend	345
	Invited to peers' houses	347
	Invites peers to own house	349
	Role play with peers	369
	Does not understand rules of polite behavior	702
	Can be tactless and inappropriate	703
	Treats everyone the same way	704
	Loner, neglected *(A1)**	716

	Odd, avoided *(A1)**	717
	***n* of items**	**15**
B1	Favorite phrases used excessively (sometimes in inappropriate situations)	695
	Uses second-hand phrases (echoes)	748
	Endless repetition of word or phrase *(B2)**	750
	Plays by making patterns rather than creatively	755
	Hand and finger manierisms	766
	Complex manierisms	767
	***n* of items**	**6**
B2	Keeps telling people things they already know	678
	Asks questions even though already knows the answers	680
	Rituals around dressing etc	756
	Resistant to changes in routes or rooms etc	758
	Resistance to change in terms of day-to-day routine	979
	***n* of items**	**5**
B3	Turns conversation to favorite theme *(B1)**	692
	Has a large store of facts	722
	Over-riding specific interests	723
	Preoccupation with something odd or bizarre that would not interest most children	754
	Preoccupation with things that might interest children only briefly	973
	Intense attachment to objects that are unusual or bizarre	975
	***n* of items**	**6**
B4	Sensitivity to everyday sounds (now)	99
	Sensitive to everyday sounds (ever)	101
	Complain music is too loud (now)	103
	Complain that music is too loud (ever)	105
	Hands over ears to everyday sounds (now)	107
	Hands over ears to everyday sounds (ever)	109
	Family adjustment to sound sensitivity	111
	Infant-like interest in smell, touch or taste of objects	757
	Unusally interested in objects that spin	977
	Studies objects from unusual angles	978
	***n* of items**	**10**
Other	Uses sophisticated or unusual words *(B3)**	721
	***n* of items**	**1**

* Categorized differently by the panel in comparison to original mapping by authors of the algorithm *(original mapping added between brackets).*

Note. A1 = Deficits in social-emotional reciprocity; A2 = Deficits in nonverbal communication; A3 = Deficits in relationships; B1 = Stereotyped/repetitive behaviors; B2 = Insistence on sameness and routines; B3 = Restricted, fixated interests; B4 = Hyper- or hyporeactivity to sensory input.

We categorized ten out of 63 items differently compared to the instrument (Mandy et al., 2017). The majority of those items ('Being regarded as rude', 'Being neglected, a loner', 'Not varying behavior according to social context', 'Less socially aware than peers', 'Inappropriate social behavior outside home') shifted from 'Deficits in social-emotional reciprocity' (A1) to 'Deficits in relationships' (A3). One A1-item was moved to A2 ('Response to parental facial expression'), and two items were reorganized under another B-sub-criterion: from B1 to B3 ('Turns conversation to favorite theme'), from B2 to B1 ('Endless repetition of words or phrases'). In addition, the current item mapping was unable to organize one item under any of the DSM-5-sub-criteria. This item – 'Using sophisticated or unusual words' – was considered too broadly formulated, and mostly associated with cognitive skills and not specifically associated with ASD and hence removed from ASD criteria, whereas it was originally categorized under 'Deficits in social-emotional reciprocity' (A1).

Table B3.

Items included in the DSM-5 algorithm of DISCO-11 were mapped onto DSM-5 sub-criteria

DSM-5 criteria	Item description		Item number
A1	Using other people as mechanical aids *(A2)**		3 x C02
	Does not seek comfort when in pain or distress		3 xii C05
	Does not give comfort to others		3 xii C06
	Makes one-sided social approaches		3 xii C07
	Lack of awareness of others' feelings *(A3)**		3 xii C08
	Does not share in others' happiness		3 xii C11 [b]
	Sharing interests limited or absent		3 xii C12
	No emotional response to age peers		3 xiii C06
	Inappropriate response to others' emotions		6 i 15
	Communication is one-sided		7 i 02
		***n* of items**	**10**
A2	Does not use nonverbal communication		3 x C01
	Does not understand gesture and miming		3 xi A01
	Lack of instrumental gestures		3 xi A05
	Lack of joint referencing pointing		3 xi A03
	Lack of imperative gestures		3 xi A02
	Lack of descriptive gestures		3 xi A06
	Does not nod or shake head		3 xi A04
	Lack of emotionally expressive gestures *(A1)**		3 xi A07
	Inappropriate or no facial expression		3 xi C05
	Non-verbal communication is absent or odd		3 xi C06
	Poor eye contact		3 xii C13
	Makes brief glances		3 xii C14
	Blank, unfocussed gaze		3 xii C15
	Stares too long and hard		3 xii C16
		***n* of items**	**14**
A3	Unusual response to visitors		3 xii C04
	No interest in age peers *(A1)**		3 xiii C01
	Does not interact with age peers		3 xiii C02
	Problems in interaction with peers		3 xiii C07
	Lack of friendship with age peers		3 xiii C09
	Poor quality of friendship		3 xiii C10 [b]
	Difficult behaviour in public places		6 i 09
	Lack of awareness of personal modesty		6 i 10 [a]
	Does not understand psychological barriers		6 i 11 [a]
	Does approach strangers inappropriately		6 i 12
	Embarrassing remarks in public		6 i 13 [b,c]
	Interrupts conversations		6 i 14 [b]
		***n* of items**	**12**
B1	Immediate echolalia		3 x C04 [c]
	Delayed echolalia		3 x C05 [c]
	Idiosyncratic use of words		3 x C07 [c]
	Long winded pedantic speech		3 x C08 [c]
	Atypical tone of voice in speech		3 xi C01 [c]
	Unusual movements of hands or arms		4 i 02
	Self-spinning		4 i 04
	Rocking (standing up)		4 i 06
	Complex movements		4 i 07
	Interests in parts of objects		4 v 05
	Elaborate repetitive activities with objects		4 v 06
	Interests in abstract properties of objects		4 v 07
	Limited pattern of activities		7 i 04
		***n* of items**	**13**
B2	Acts out role of figure / person repetitively		3 xv C05
	Arranges objects in patterns		4 v 04
	Insists on sameness of environment		4 v 08
	Insists on perfection		4 v 09
	Has unusual food fads		4 v 10
	Insists on sameness in routines		4 v 11
	Clings to home or other familiar places		4 v 12

	Has other repetitive routines	4 v 19
	Asks repetitive questions	4 v 13 [c]
	Talks about a repetitive theme	4 v 14 [c]
	***n* of items**	**10**
B3	Clings to objects	4 v 01
	Collects objects	4 v 02
	Fascinated with specific objects	4 v 03
	Talks about a repetitive theme *(B2)**	4 v 14 [c]
	Has repetitive activities related to a special skill	4 v 15
	Collects facts on specific subjects	4 v 16
	Fascinated with TV/videos	4 v 17
	***n* of items**	**7**
B4	Smells objects or people	4 ii 05
	Unusual interest in the feel of surfaces	4 ii 06
	Repetitive, aimless manipulations of objects (not near eyes) as if seeking sensory stimulation	4 ii 09
	Indifference to pain, heat or cold	4 ii 11
	Distress caused by sounds	4 iii 01
	Fascinated with sounds	4 iii 02
	Fascinated with bright lights and shiny objects	4 iv 01
	Interest in watching things spin	4 iv 02
	Twists hand or objects near eyes	4 iv 03
	Studies the angles of objects	4 iv 04
	***n* of items**	**10**
C	Setback in social development	3 i 2f
	Delay in selective social attachment	3 xii B01
	***n* of items**	**2**
Other	*Setback in language development (C)**	*3 i 2d*
	*Delay in obeying instructions (C)**	*3 ix B01*
	*Age started to use phrases (C)**	*3 x B02*
	*Content of speech difficult to understand (B2)**	*3 x C10* [c]
	*Lack of pretend play (A3)**	*3 xv A02*
	*Delay in the development of pretend play (C)**	*3 xv B02*
	*Shows anger toward parents (A3)**	*6 i 06*

* placed differently in comparison to original mapping by authors *(original mapping added between brackets)*

[a] not applicable for children younger than 4y, [b] not applicable for children younger than 6/7y, [c] not applicable for minimally verbal individuals

Note. A1 = Deficits in social-emotional reciprocity; A2 = Deficits in nonverbal communication; A3 = Deficits in relationships; B1 = Stereotyped/repetitive behaviors; B2 = Insistence on sameness and routines; B3 = Restricted, fixated interests; B4 = Hyper- or hyporeactivity to sensory input; Criterion C = Early onset.

Some disagreement was found when we compared the current mapping with the original organization of items by the authors (see Table B3). Two items of A1 were moved to A2 ('Lack of emotionally expressive gestures') and A3 ('No interest in age peers') respectively. One item of A3 was placed within A1 ('Lack of awareness of others' feelings'), and one of A2 was placed under A1 ('Using other people as mechanical aids') and one of B2 within B3 ('Talks about a repetitive theme'). Another seven items were coded as not belonging to any of the DSM-5criteria, because they were either not specifically associated to ASD and related to behavior problems in general (item 'Anger towards parents', originally A3), or because they refer to behaviors that are no longer part of ASD symptoms anymore (items 'Lack of pretend play' (A3), 'Content of speech' (B2), 'Setback in language development' (C), 'Age started to use phrases' (C), 'Delay in development of pretend play' (C) and 'Delay in obeying instructions' (C)). In particular the items for criterion C are based on DSM-IV-TR and not up-to-date with

the current DSM-5 description. The authors explicitly state that for Criterion C no changes were made compared to the previous version, as was commonly done for other papers at that time.

Appendix C. Subscales included in the 3di algorithm

Table C1.

Subscales included in the DSM-5 algorithm of 3di [6]

DSM-5 criteria	**3di**
A - Deficits in social communication and interaction	**3/3 sub-criteria**
A1 - Deficits in social-emotional reciprocity	**1/5 subscales:**
	A1a. Number of abnormal social approach symptoms
	A1b. Age appropriate social behavior
	A1c. Failure of back and forth conversation
	A1d. Sharing
	A1e. Sharing of emotions/reciprocity
A2 - Deficits in nonverbal communication	**1/4 subscales:**
	A2a. Poorly integrated verbal and non-verbal
	A2b. Eye contact and body language
	A2c. Facial expression
	A2d. Gestures
A3 - Deficits in relationships	**1/3 subscales:**
	A3a. Adjusting behavior to social context
	A3b. Sharing imagination
	A3c. Friendships
B - Restricted, repetitive behavior, interests or activities	**2/4 sub-criteria**
B1 - Stereotyped / repetitive behaviors	**1/2 subscales:**
	B1a. Repetitive behavior
	B1b. Repetitive speech
B2 - Insistence on sameness and routines	**1/2 subscales:**
	B2a. Verbal rituals and routines
	B2b. Nonverbal routines and rituals
B3 - Restricted, fixated interests	**1/2 subscales:**
	B3a. Fixated on objects
	B3b. Focused interests
B4 - Hyper- or hyporeactivity to sensory input	**1/2 subscales:**
	B4a. Hyposensitivity
	B4b. Hypersensitivity

References

1. American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders (5th edition). Author, Washington, D.C.

2. Lord C, Luyster R, Gotham K, Guthrie W (2012) Autism Diagnostic Observation Schedule, second edition (ADOS-2): Manual. Part II: Toddler Module. Western Psychological Services, Torrance, CA

3. Lord C, Luyster R, Gotham K, Guthrie W (2012) Autism Diagnostic Observation Schedule, second edition (ADOS-2): Manual. Part I: Modules 1-4. Western Psychological Services, Torrance, CA

4. Kent R, Carrington S, Couteur A, Gould J, Wing L, Maljaars J, Noens I, Berckelaer-Onnes I, R Leekam S (2013) Diagnosing Autism Spectrum Disorder: who will get a DSM-5 diagnosis? J Child Psychol Psychiatry 54:1242–1250

5. Carrington S, Kent R, Maljaars J, Le Couteur A, Gould J, Wing L, Noens I, Van Berckelaer-Onnes I, Leekam SR (2014) DSM-5 Autism Spectrum Disorder: In search of essential behaviours for diagnosis. Res Autism Spectr Disord 8:701–715 . doi: 10.1016/j.rasd.2014.03.017

6. Mandy W, Wang A, Lee I, Skuse D (2017) Evaluating social (pragmatic) communication disorder. J Child Psychol Psychiatry 58:1166–1175 . doi: 10.1111/jcpp.12785